Automated knot detection with visual post-processing of Douglas-fir veneer images

C.L. Todoroki [a,*], E.C. Lowell [b], D. Dykstra [b]

[a] Scion, Private Bag 3020, Rotorua 3046, New Zealand
[b] USDA Forest Service, Pacific Northwest Research Station, P.O. Box 3890, Portland, OR 97208, USA

ARTICLE INFO

ABSTRACT

Knots on digital images of 51 full veneer sheets, obtained from nine peeler blocks crosscut from two 35-foot (10.7 m) long logs and one 18-foot (5.5 m) log from a single Douglas-fir tree, were detected using a two-phase algorithm. The algorithm was developed using one image, the Development Sheet, refined on five other images, the Training Sheets, and then applied to all remaining sheets. In phase one, global thresholding was used to segment the image through a series of morphological operations to isolate regions likely to contain knots. In phase two, adaptive thresholding was applied to grey scale and red component segmented images to improve the accuracy of the segmented knot.

Overall performance, judged in terms of confusion matrix performance metrics, was better for the red component images. Red component recall (true positive) rate was 1.00, 0.99, and 0.96 for the Development, Training, and complete sets, respectively. For the grey scale images, recall rates were 0.96 for all sets. Red component accuracy was 0.76, 0.92, 0.73 (Development, Training, and complete) and those for the grey scale images were 0.71, 0.85, and 0.69, respectively. Red component precision also exceeded that of the grey scale (0.75, 0.93, 0.73 compared to 0.72, 0.88, 0.70). A greater percentage of knots (78%) segmented from red component images were correctly sized, while 16% had more pixels than required and 6% had fewer pixels. Comparative figures for the grey scale images were 57% correctly sized, 2% with more pixels, and 42% with less pixels.

Based on our results, we will adopt the red component image for continuing work with digital veneer images from a sample of Douglas-fir trees selected on the basis of acoustic velocity measures. Together with acoustic measurements of the veneer sheets, we are investigating the extent that the number, size, and spatial arrangement of knots influences the average stiffness of veneer sheets, with a view to determining if a relationship exists between the average stiffness of veneer sheets in a peeler block, stiffness of the log, and stiffness of the parent tree from a range of silvicultural treatments.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

1.1. Stiffness and acoustic measurements

Acoustic technologies for assessing stiffness, such as Fibre-Gen's Director HM200™, FAKOPP's TreeSonic, and the Metriguard 2600™ (a product of Metriguard Inc.) have become increasingly popular in forest and processing environments. Their popularity arises from the tools being relatively inexpensive, simple to use, and because they permit testing of wood samples to be done non-destructively. This provides opportunities for better resource quality assessments (Chauhan and Walker, 2006; Cown, 2005), better log segregation into quality classes (Dickson et al., 2004; Amishev and Murphy, 2008), better board segregation into stiffness classes (Carter et al., 2006), and early screening for genetic heritability (Kumar et al., 2002), yielding great potential to add value all along the forest-to-products chain. The earlier well-informed decisions are made within the forest-to-products chain, the greater the potential value addition. Thus if it can be shown that strong relationships exist between tree and product, then those relationships can then be used to generate added value.

Relationships between trees and products have been investigated by Dickson et al. (2004), Lasserre et al. (2007), Raymond et al. (2008), and others, for radiata pine (Pinus radiata D Don). Dickson et al. (2004) found acoustic velocity measurements of standing trees sourced from thinned (350 stems/ha) plots to be closely correlated to that of the logs cut from those trees. They also found the relationship between velocity measurements of those logs and machine stress grades of the constituent boards to be even stronger, particularly for boards sawn from outerwood. Lasserre et al. (2007)

* Corresponding author. Tel.: +64 7 3435899; fax: +64 7 3435507.
E-mail addresses: christine.todoroki@scionresearch.com
(C.L. Todoroki), elowell@fs.fed.us (E.C. Lowell), ddykstra@fs.fed.us (D. Dykstra).

demonstrated that branches and bark influence acoustic velocity, and when stems are delimbed and the bark is removed, velocity, and hence stiffness, significantly increases. Stiffness decreases linearly up a stem and has been found to be lower at each height in the stem for trees grown on thinned sites (Raymond et al., 2008). For Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco), acoustic velocity readings also decrease with increasing height in the stem, and demonstrate good correlation, when combined with green density measurements, with veneer grade recovery (Amishev and Murphy, 2008).

In addition to the presence of branches providing a source of variation for acoustic measurements, knot size also contributes to the variation. Amishev and Murphy (2008) found both the presence and size of knots to be negatively correlated to acoustic velocity and suggested that acoustic measurements could be incorporated into optimal crosscutting tools (known also as bucking tools) based on market requirements for wood stiffness by accounting for the effect of branch size and other variables.

1.2. Knot detection and performance measures

To add to the sources of variation, knots and other defects come in a variety of sizes, shapes, types, and colours. When boards are manually inspected, the ability to accurately recognise, locate and identify surface knots has been assessed at 55% (Polzleitner and Schwingshakl, 1992) and at 68% (Huber et al., 1985). Even with automated systems for detecting knots in hardwood lumber, Buehlmann et al. (2007) suggest that a 100% accuracy rate is not always achievable.

In research environments, accuracy rates exceeding 70% and approaching the 100% level have been achieved. Using a two stage approach known as sweep-and-mark (Forrer et al., 1988), defect accuracies on Douglas-fir veneer images of 94, 89, and 79% were achieved with statistical, morphological, and colour-cluster algorithms, respectively. When extending their research to exclude pitch pockets Forrer et al. (1989) found that accuracies increased to 99, 98, and 86%, respectively.

In addition to accurately detecting the presence or absence of knots and other defects on Douglas-fir veneer sheet images, Funck et al. (2003) examined the accuracy of defect size accuracy, measured on a pixel basis. Overall, pixel accuracy of loose and tight knots averaged about 95 and 90%, respectively. The accuracy was not significantly different across a range of colour spaces obtained from Douglas-fir veneer images, nor has it been found to differ significantly for transformed colour spaces (Brunner et al., 1992). However, with image segmentation performed via a fuzzy neural network on radiata pine boards, Ruz et al. (2005) achieved a pixel accuracy of 94%, along with a 95% defect detection rate and a 6% false positive rate.

Accuracy and other performance metrics can be derived from a confusion matrix (Kohavi and Provost, 1998; Fawcett, 2006). The 2-by-2 matrix contains four cells: true positives (TPs), false positives (FP), true negatives (TNs), and false negatives (FN). In some settings, the correct identification of TPs is important, whereas in other settings it may be more important to correctly identify TNs (Sokolova and Lapalme, 2009). For our research, the true positive rates and false negative rates are of primary importance. The true positive rate, also known as the recall rate, is the proportion of positive cases that are correctly identified (Eq. (1)). The false negative rate is the proportion of positives cases that are incorrectly classified as negative (Eq. (2)).

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{1}$$

$$\text{FN rate} = \frac{\text{FN}}{(\text{TP} + \text{FN})} \tag{2}$$

Other performance measures derived from the confusion matrix include accuracy and precision. Accuracy is the proportion of the total number of predictions that are correct (Eq. (3)) while precision is the proportion of positively identified cases that are correct (Eq. (4)).

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})} \tag{3}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{4}$$

1.3. Veneer stiffness project

In our research, unlike other studies that consider the part of the forest-to-products chain from logs to lumber (e.g. Abbott and Kline, 2003; Rinnhofer et al., 2003), we consider the chain from tree to veneer. Throughout that chain, we examine stiffness via acoustic measurements. Our research extends to examining knots on the acoustically measured veneer sheets. This will enable the development and evaluation of models that include not only acoustic but also knot variables for improved prediction of stiffness. This paper focuses on the knot detection portion of that larger project in a desktop application. The aim of our knot detection project is to accurately detect knot location and size on digital veneer images with particular emphasis on achieving high recall rates. The overall aim of the larger project is to determine if a relationship exists between the average stiffness of veneer sheets in a peeler block, stiffness of the log, and stiffness of the parent tree, and to investigate the extent these relationships are influenced by silvicultural treatment, stand, tree, log, and knot variables.

2. Method

2.1. Sample material

The Stand Management Cooperative (SMC) managed by the University of Washington and involving a large number of forest landowners and other co-operators, has an extensive system of Douglas-fir research installations throughout the coastal Pacific Northwest. Installations were established between 1987 and 1989. At the time of establishment, five permanent plots (a control plot plus four treatment plots that followed prescribed thinning regimes) were set up at each installation. For this project, four SMC installations representing four different ownerships were selected; two in Washington and two in Oregon. One of the installations had only four plots as the fifth, a treatment plot, had storm damage. Project logistics required these installations to be reasonably accessible by road.

A circular 0.25 acre (0.1 ha) plot was established in the geometric centre of each of the five plots. Time-of-flight acoustic speed was obtained over a 1 m distance approximately centred at breast height (with one probe placed about 50 cm below breast height and the other probe placed approximately the same distance above breast height) on all of the trees within the circular plot using an acoustic device. Three readings were obtained at each of three locations approximately 120° apart around the stem circumference, providing nine observations per tree. The nine stress wave times were converted to mean acoustic velocity for each tree. Trees on each plot were listed in ascending order of acoustic velocity and a stratified random sample was selected. Two trees were randomly chosen from the lowest 10%, four from the next 11 to 50%, four from 51 to 90%, and two from the top 91 to 100% on each plot. One half of the trees in each stratum were randomly chosen for veneer conversion and the remainder for lumber conversion (Briggs et al., 2008).

Table 1
Summary of data for the parent tree, logs, peeler blocks, and veneer sheets used in the image analysis.

SMC installation	Plot	Tree	Diameter at breast height (in.)	Total height (ft)	Mean acoustic time-of-flight (ms/m)	Acoustic class	Long log	Peeler block	Large end diameter (in.)	Small end diameter (in.)	No. of sapwood half sheets	No. of full sheets	No. of heartwood half sheets
803	1	124	11.9	108	258	3	1	1	13.6	11.4	0	9	0
								2	11.4	10.8	0	8	1
								3	10.8	10.7	0	8	0
								4	10.7	10.4	0	8	0
							2	1	10.4	10.4	0	6	0
								2	10.4	9.2	0	5	2
								3	9.2	8.4	1	3	0
								4	8.4	7.2	0	3	0
							3	1	7.2	6.2	0	1	0
								2	6.2	5.0	–	–	–

Notes: Acoustic class 3 comprises trees within the plot that had mean acoustic velocities between 51 and 90% of the maximum acoustic velocity measured on the plot. Peeler block 3–2 was discarded at the mill and not peeled.

Trees were felled, delimbed, and topped at the merchantable diameter limit of 5 in. (12.7 cm). The acoustic velocity in each of the resulting delimbed stems was measured with an acoustic device. Veneer sample trees were next crosscut into long logs (target lengths of 35 ft (10.7 m) including trim) and again tested acoustically. Long logs were further crosscut into mill-length logs (17.5 ft (5.3 m) including trim) and re-tested with the acoustic device. Mill-length logs were crosscut into two veneer blocks, each 8 ft 4 in. (2.5 m) long. This latter operation was done at the mill and the blocks were then individually measured and re-labelled for identification purposes. No acoustic testing was done on the veneer blocks. The 697 blocks in the study sample were steamed overnight and then peeled into veneer over a period of 2 days. During the peeling process, a tracking methodology was used that allowed each sheet to be traced back to the peeler block and consequently to the position within the parent tree from which it came. Each veneer sheet from a particular block was labelled with the block number and the order in which it was peeled from the block. Veneer thickness was 0.15 in. (3.8 mm) and the target peeler core diameter was 3.5 in. (89 mm). Full veneer sheets had dry target sizes of 102 in. (2.59 m) in length and 52 in. (1.32 m) in width. After being dried at the veneer mill, the veneer sheets were shipped to the US Forest Products Laboratory (FPL) in Madison, WI where they were non-destructively tested for stiffness using a Metriguard 2600™ veneer tester.

2.2. Veneer images

The photographic analysis described in this paper is based on high-resolution digital images (3872 × 2592 pixels, 24-bit RGB colour depth) taken by FPL technicians using a Nikon D80 camera mounted about 10 ft (3 m) directly above the veneer sheet. For lighting, the camera's flash was turned off and two 600 W quartz lamps were used, mounted about 8 ft (2.4 m) above the veneer sheet at angles of 45°. The sheet being photographed was placed on a frame with measuring tapes affixed to the four sides of the frame for reference.

The images selected for this analysis represent all the full veneer sheets (51 full sheets that came from nine peeler blocks) produced from a single Douglas-fir tree (number 124 from treatment plot 1 on SMC installation 803). The installation is located on flat ground in private ownership near the small community of Matlock, in western Washington. Data on the parent tree, 9 peeler blocks, and 51 veneer sheets produced are summarised in Table 1. A tenth peeler block, the topmost block cut from the stem, was rejected at the mill due to its small diameter, large knots, and irregular form. It was sent to the chipper rather than the peeler and therefore is not included in this analysis.

In addition to the 51 full veneer sheets, 4 half sheets were peeled from this tree. As far as we have been able to determine these

Fig. 1. Original image.

comprise all of the usable veneer that was peeled from this tree. Information on roundup scraps and broken sheets was not collected, so some of the wood from the nine peeled blocks has not been accounted for.

2.3. Digital image processing

Algorithms to detect knots on grey scale and red component veneer images were developed using MATLAB® and its image processing toolbox (Gonzalez et al., 2004). The RGB images were pre-processed to reduce the raw image (Fig. 1) to the area of interest, realign the image to the viewing window, and to inscribe a rectangle within the tape-measure to represent the veneer sheet boundaries (Fig. 2). Knot detection algorithms worked within that bounded region and were developed using the veneer sheet shown
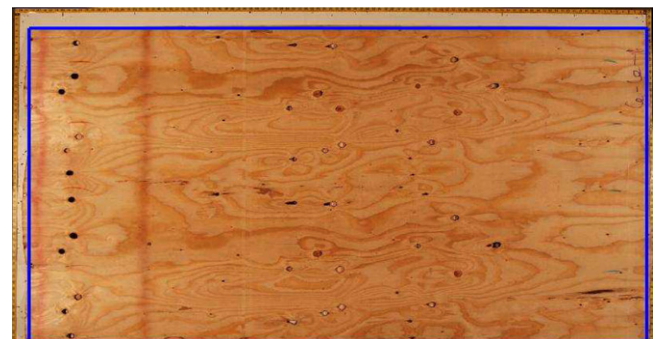
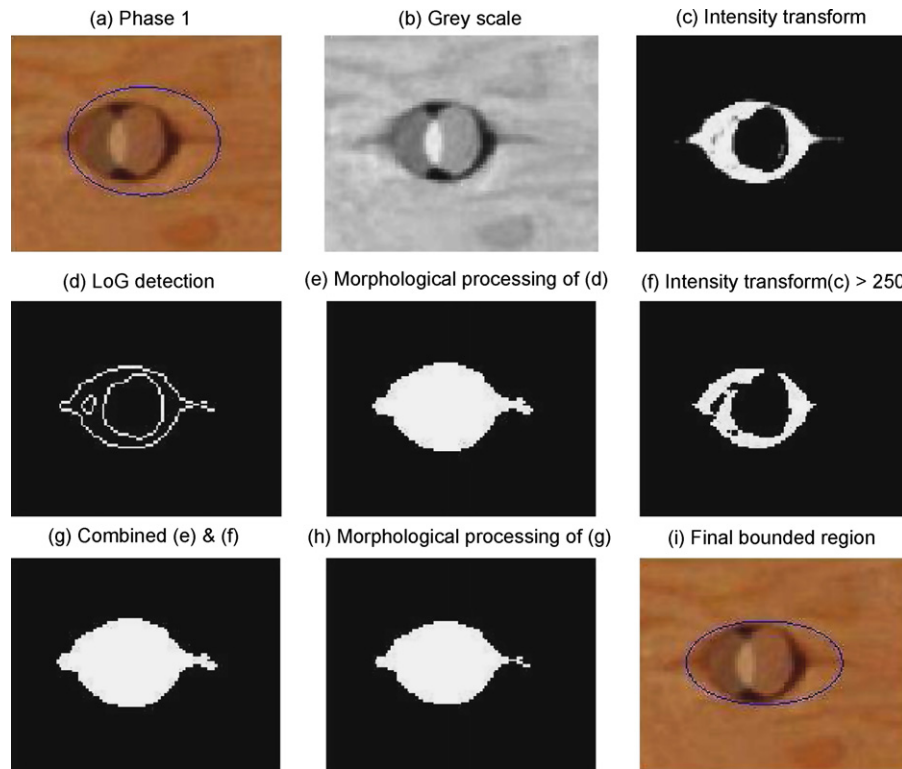**Fig. 2.** Bounded veneer sheet used in image processing.

Fig. 3. Steps of the knot detection applied to a grey scale image.

in the figure (known hereafter as the Development Sheet), and tuned using five other veneers sheets (the Training Sheets). The Development Sheet was the first sheet of the first peeler block of Long Log 2, thus was from the block in the middle of the tree (Table 1). It was chosen simply as it was the first sheet in the veneer labelling sequence. The Training Sheets were the next five consecutive sheets peeled from that block. After the knot detection algorithms had been tuned, they were applied to the remaining veneer sheets. All 51 images of full sheets, from 9 blocks from the same tree, were processed in this way.

The knot detection algorithms comprised two phases. In phase one a global approach was applied to the veneer image, and after conversion to grey scale and a series of morphological operations, regions that demonstrated potential to contain knots were segmented. In phase two, a local approach was applied to each of the segmented regions. The purpose of the phase two algorithm was to refine and improve on the accuracy of the knot's location and size.

The phase one global approach was applied to a grey scale image. Global threshold values, which differed from sheet to sheet, were assigned automatically using Otsu's method (Otsu, 1979). Scaled threshold values were then used to transform the intensity of the veneer image to enhance the "knotty" regions. The intensity transformation was achieved by mapping the intensity values between 50 and 75% of the threshold value, T, obtained from Otsu's method, thus clipping out values less than 50% of T and greater than 75% of T. The values within the defined range $[0.50T, 0.75T]$ were mapped to the full $[0, 1]$ range. A binary image was then formed from the intensity transformed image using a relational operator with a threshold value of 50. Morphological operations that first dilated the binary image using a rectangular structuring element, then eroded the resultant image using a disk-shaped structuring element created objects that were subsequently traced and labelled as knots. In several cases, features other than knots, i.e. false positives, were segmented by this process.

From a practical viewpoint, false positives are not as much of a problem as false negatives. This is because removing a false positive reading is faster and easier than inserting a region of correct proportions to correct a false negative. Thus the false positive rate was not a concern and the effectiveness of the algorithms was judged by the true positive and false negative rates. The aim was to develop an algorithm that maximised the true positive rate and hence, as the two rates sum to one, minimised the false negative rate.

The phase two refining algorithm worked with both grey scale and red component images. The main steps within phase two are described below and illustrated by way of example in Fig. 3(a)–(i) (grey scale) and Fig. 4(a)–(i) (red component).

Each region delineated in phase 1 was cropped around the phase 1 solution (Figs. 3a and 4a). The cropped image is not only smaller in size than the original image, but also requires less memory to process, and allows faster computations. A grey scale (Fig. 3b) or red component (Fig. 4b) image was formed from the cropped image, then subjected to an intensity transformation to enhance the knotty area (Figs. 3c and 4c). This was achieved by application of Otsu's method to retrieve a new threshold value for the cropped image. The threshold, Tc, was then used in an intensity transformation that scaled values between [0.9Tc, 0.9Tc + 0.05] to the full range. The complement of the intensity transformed image was then computed and a binary image formed by thresholding with a value of 250 using a relational operator. Subsequent segmentation of the image through application of the Laplacian of a Gaussian (LoG) detector, which takes the second derivative (the Laplacian) of the Gaussian function $h(r)$, given by

$$h(r) = -e^{-(r^2/2\sigma^2)}$$

where $r^2 = x^2 + y^2$ and σ is the standard deviation, isolated the edges (Figs. 3d and 4d). A combination of morphological image processing techniques, dilation and erosion using a disk-shaped
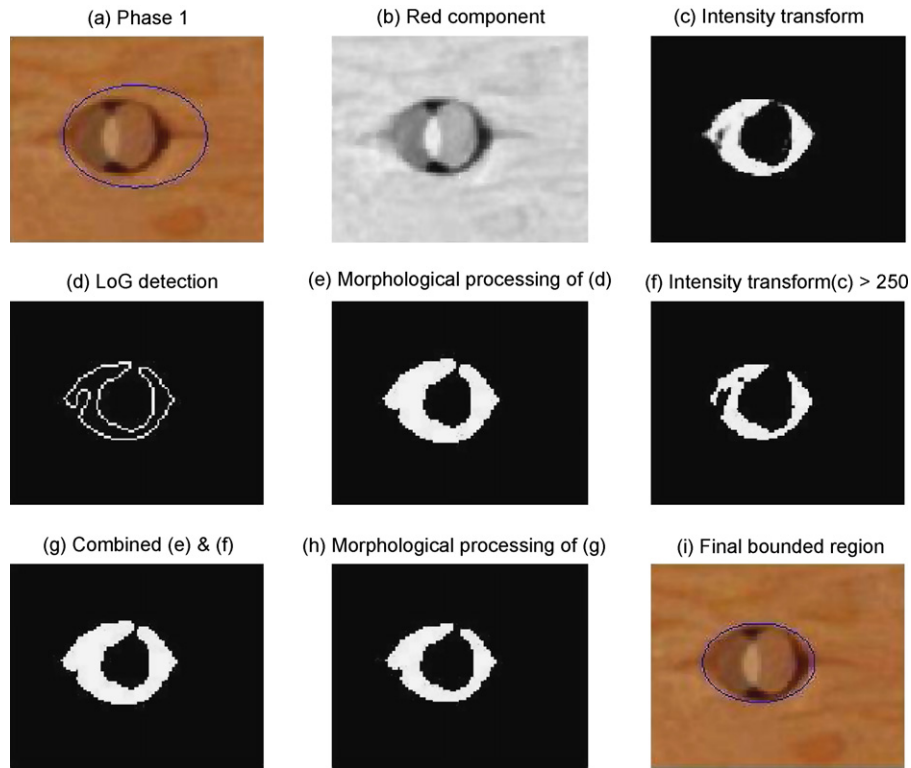
Fig. 4. Steps of the knot detection applied to a red component image.

structuring element, were then applied and any holes were filled (Figs. 3e and 4e). In our testing of the phase two algorithm with a range of knots from the Development sheets, we found that some parts of knots were not picked up by the edge detection method. For this reason a Boolean image (Figs. 3f and 4f) that was created by thresholding (using a 250 threshold) the intensity transformed image (Figs. 3c and 4c) was combined with the filled image using a logical operator (Figs. 3g and 4g). A further series of morphological operations using a disk-shaped structuring element were performed, giving the image components shown in Figs. 3h and 4h. Bounded regions were then created by tracing the existing objects, to determine the final knot size and location (Figs. 3i and 4i).

In cases where more than one bounded region was detected, the region most likely to contain a knot was determined by filtering out less likely regions and prioritising the remaining regions. Less likely regions were defined as those for which the height exceeded the width by more than 30%, those for which the height exceeded 1.5 in., and those for which the height and width were each less than one-eight of an inch. Knots were prioritised according to distance and size. Distance, d, was measured as the Euclidean distance between the centres of the original segmented region (x_c, y_c) and that region under investigation (x_r, y_r).

That is, $d = [(x_c - x_r)^2 + (y_c - y_r)^2]^{1/2}$

Size was measured in terms of the region's area, A. The ratio of area over the square of distance was then computed and the maximum, p, selected, thus favouring larger regions in close proximity to the original region.

i.e. $p = \max[A_i / d_i^2]$

for all bounded regions i within the cropped image

After the algorithms had run their course and the segmented regions (comprising both true positives and false positives) had been plotted on the veneer sheet image, false negatives (i.e. those knots that should have been detected but were not) were determined by visual inspection.

2.4. Algorithm performance measures

Confusion matrix entries (Table 2) were counted for each of the grey scale and red component approaches for each veneer sheet. The two approaches were compared, on a sheet by sheet basis, in terms of the classifier performance indicators: recall and false negative rates, accuracy, and precision, as calculated from Eqs. (1), (2), (3), and (4), respectively.

Further performance measures, size accuracy and preference, were introduced to indicate the accuracy in size of the segmented regions, and to indicate which segmented region (grey or red) provided the more accurate representation and hence was the preferred option. These measures were visually assessed.

Size accuracy was assessed for the true positives, i.e. those knots that were correctly detected and measured in terms of three variables: accurate, under-sized, or over-sized. "Under-sized" indicated that insufficient pixels had been captured during the segmentation process. The contrary was true for "over-sized". The choice of these variables, rather than a pixel count, was based on the premise that no matter the extent to which a segmented region surrounding a knot is over-sized or under-sized, it still needs a series of operations (e.g. delete region, and manually drag and drop a new region) to produce an accurately sized region. The distinction was made between under-sized and over-sized to observe any differences between the two (grey vs. red) approaches.

Image component preference indicated which of the two images achieved the better fit to actual knot size and location. Preference was measured using Boolean variables; (1, 0) indicated that the knot segmented using the grey scale image was preferred. The converse was true for (0, 1). And in cases where the results appeared

Table 2
Confusion matrix with cell definitions.

	Actual	
Detected	p	n
Y	True positive	False positive
	Actual knot correctly classified	Actual feature, other than a knot, incorrectly classified as a knot
N	False negative	True negative
	Actual knot not detected and hence not classified	Actual feature, other than a knot, detected in phase 1 and correctly removed from analysis in phase 2
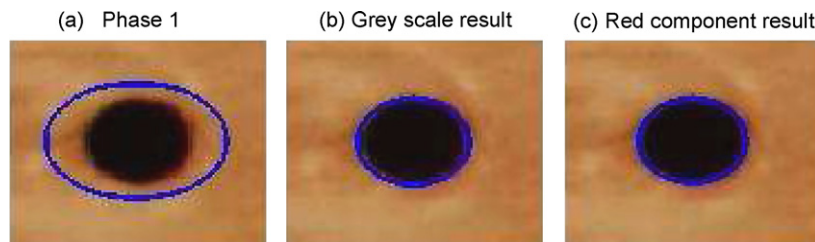
Fig. 5. Accurately detected grey scale and red component knots.
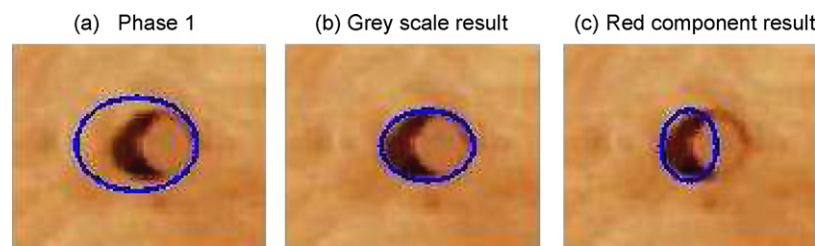
Fig. 6. Over-sized (grey) and under-sized (red) knot detection.

identical (after zooming in) each method recorded a preference indicator (1, 1).

For example, while both segmented regions in Figs. 3i and 4i are over-sized, the red component image clearly demonstrates a better fit, hence is preferred. Fig. 5 illustrates a knot where both grey and red knot images are assigned an "accurate' status for region size. The grey scale image represents the outer limit of what passed as "accurate" with some lighter coloured pixels visible to the left of the knot. Modelling preference was allocated to the red scale image. In Fig. 6, the grey scale region is "over-sized", the red component region "under-sized" and the preference allocated to the grey scale image.

Interpretation of the performance metrics was achieved through graphical inspection of results combined with regression and statistical analysis. Excel®'s TTEST statistic was used to determine whether the samples (recall rate, accuracy, precision, and preference) from the grey scale and red component images were likely to have come from the same two underlying populations having the same mean. The Excel function output is the fractional probability of the Student's t-distribution. Thus a returned value of 0.05 corresponds to a 95% confidence level for rejecting the null hypothesis. The null hypothesis is that there is no difference between the two means.

3. Results

3.1. Confusion matrix performance metrics

True positive (recall) rates for the complete sets of red component and grey scale images were not significantly different. However, there was a clear, and statistically significant, difference in accuracy and precision, both of which were greater for the red

component images. Knot size accuracy was also greater with the red component images leading to an average 90% preference score for red component segmented knots.

3.2. Recall rate (Eq. (1))

The confusion matrix for the Development Sheet shown in Table 3 contains entries for both grey scale and red component images. The veneer sheet contained 113 knots (TP + FN) of which 99 were accurately detected using the grey scale image and all 113 detected with the red component image. Hence the recall rate for grey scale and red component Development Sheet images was 0.96 and 1.00, respectively.

Mean recall rate with the Training Sheets was 0.96 for the grey scale images (ranging from 0.942 to 0.975) and 0.99 (ranging from 0.975 to 1.00) for the red component images, each with standard deviations of 0.01. For the complete set of images, mean recall rate for both grey scale and red component images was 0.96 with standard deviations of 0.02 and 0.03, respectively. Differences between means, as measured by a paired student's t-test, were not significant ($p = 0.35$). Recall rates obtained for the 51 veneer sheets are shown in Fig. 7. The Development and Training Sheets are shown as black squares.

Table 3
Confusion matrix entries for knots detection on the grey scale and red component images of the Development Sheet.

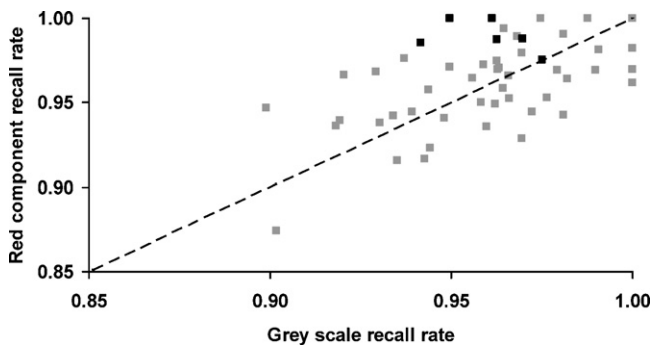|	p		n	
	Grey	Red	Grey	Red
Y	99	103	38	34
N	4	0	2	4

Fig. 7. Recall rates, grey scale vs. red component, rates of the 51 veneer sheets.

In many cases recall rates of the testing sheets exceeded the lower bounds of the grey and red component Training Sheets (i.e. were in excess of 0.942 for grey and 0.975 for red). Of the 46 testing sheets, 34 grey and 10 red sheets exceeded the lower bounds. Furthermore 12 of those grey sheets had recall rates that exceeded the upper bounds of the Training sheets (i.e. in excess of 0.975). Three red testing sheets attained the upper red bound with perfect recall rates of 1.0.

3.3. False negative rate (Eq. (2))

The false negative rate, i.e. ratio of actual knots that were not detected over the total number of knots on each veneer sheet, averaged 0.04 (standard deviation 0.03) for both grey scale and red component images. This corresponds to an average of 5 knots per sheet.

3.4. Accuracy (Eq. (3))

Accuracy, in comparison to the recall rates, was lower for both grey and red component images. This was due to false positive readings. Accuracy associated with the grey scale Development Sheet image was 0.71 while that for the red component image was 0.76. For the Training Sheets, mean accuracy increased to 0.85 and 0.92, with standard deviations of 0.01 and 0.03, for grey and red images, respectively. The greater accuracy of the Training Sheets is clearly seen in Fig. 8 with the cluster of black boxes at the top right of the figure. Accuracy between the means of the grey scale and red component images was significantly different, averaging 0.69 (standard deviation 0.12) and 0.73 (standard deviation 0.13) for grey scale and red component images, respectively. The increase in accuracy due to the red component image was equal to about 5%.
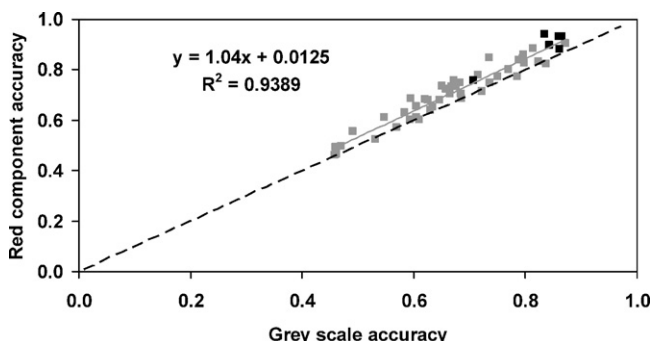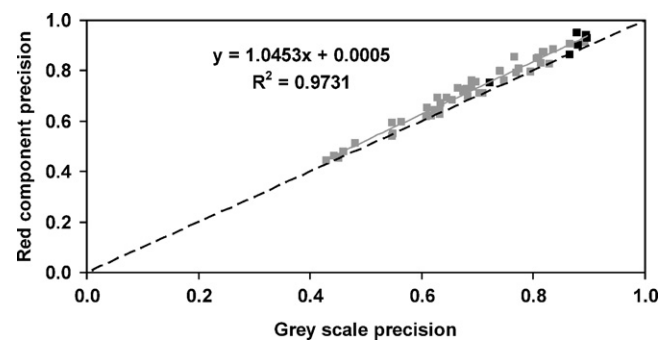
3.5. Precision (Eq. (4))

Like accuracy, precision was also influenced by false positive readings. Precision associated with the grey scale Development Sheet image was 0.72 while that for the red component image was 0.75, respectively (indicated by the black box at (0.72, 0.75) in Fig. 9). For the Training Sheets, mean precision increased to 0.88 for the grey scale images and to 0.93 for the red component images, with standard deviations of 0.01 and 0.02, respectively. The five associated points are shown by the black boxes clustered in the top right corner of the figure. As all points in the figure lie on or above the diagonal, the greater precision due to the red component image is clearly evident. Overall, mean precision was 0.70 and 0.73 for grey and red images, respectively. Furthermore, the paired t-test statistic ($p < 10^{-10}$) strongly indicated that the difference between the means (0.70 and 0.73 with standard deviations of 0.13 and 0.14 for grey and red, respectively) was highly significant. The difference, i.e. increase in precision, associated with the red component image was equal to approximately 5%.

3.6. Size accuracy

On average, only 57% of correctly segmented grey scale knots accurately represented actual knot size. For the red component images, 78% were accurately sized. The difference between the two means was statistically significant ($p < 10^{-15}$). A greater percentage of under-sized segmented knots were associated with the red component images (averaging 2% grey, 6% red), while a greater percentage of over-sized segmented areas were associated with the grey scale images (42% grey, 16% red), Fig. 10.

3.7. Preference

There was a much stronger preference for knots correctly detected through segmentation of the red component image than

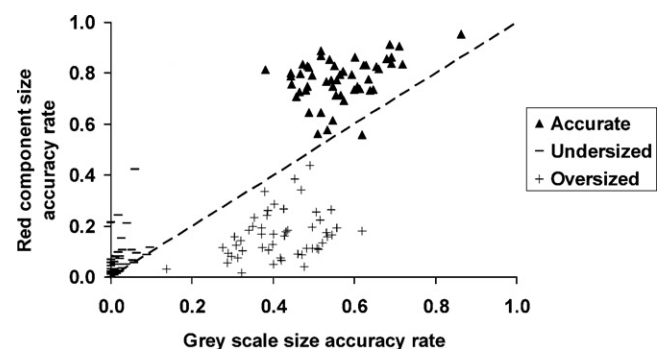**Fig. 9.** Mean precision, grey scale vs. red component, of the 51 veneer sheets.

Fig. 8. Mean accuracy, grey scale vs. red component, of the 51 veneer sheets.

Fig. 10. Mean knot size accuracy, grey scale vs. red component, of the 51 veneer sheets.
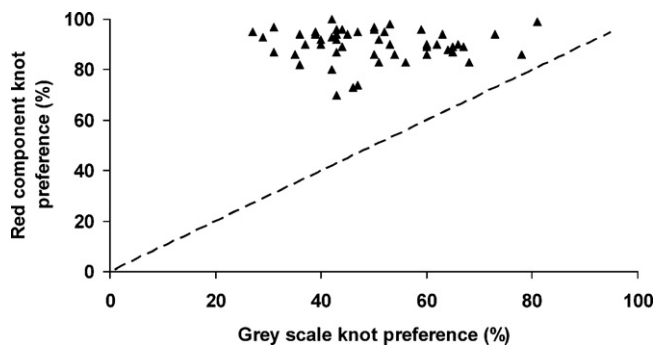
Fig. 11. Mean knot preference, grey scale vs. red component, of the 51 veneer sheets.

from grey scale images (Fig. 11). Mean modelling preference for knots detected from red component images was 90% while that for grey scale images was 50%. The difference between means was significant ($p < 10^{-15}$).

4. Discussion

The aim of our knot detection project was to accurately detect knot location and size on digital veneer images with emphasis on achieving high recall rates. Mean recall rate for both grey scale and red component images was 0.96, with Training Sheets recording means of 0.96 and 0.99, respectively. These results indicate that the algorithm performed well within expectations. While there was no significant difference between mean recall rates for grey scale and red component images, knot size accuracy was significantly different. Knots on the red component images were more accurately sized, with 78% being perfectly bounded.

In terms of accuracy and precision, the red component images out-performed grey scale images by about 4–5%. As red component images also gained considerably greater preference scores we have adopted the algorithms utilising red component images for the larger veneer project.

Due to false positive readings, rates for accuracy and precision were much lower than the recall rates. However we were not concerned with false positives, as these could be easily removed in post-processing to increase our performance metrics to 100%.

The requirement in our subsequent research for a 100% knot detection rate, which is generally not possible with machine vision applications (Buehlmann et al., 2007), meant that some post-processing work remained to perfect the data capture. In this desktop application using MATLAB® we used a mouse-click to remove the false positive regions. To reduce the false negative rate to zero we added, through interactive techniques, about 5 knots to each sheet. To perfect knot size we reduced the size of about 16% of regions and increased the size of about 6% of regions. Those regions requiring reductions were frequently associated with very small knots of less than 0.25 in. that had been over-estimated (often by only a few pixels) as a consequence of the morphological dilation operations. One way in which we considered tackling this problem was to adapt the size of the structuring element to the area of the segmented element. However, as the automated knot detection algorithms presented here have done the bulk of the otherwise time-consuming work, we could achieve the 100% level of accuracy with a minimal amount of post-processing.

In addition to the 51 full veneer sheets obtained from the single tree used here, 4 half sheets were obtained. Trials with these suggest that the algorithm works equally effectively with these and that no additional changes are needed.

We next intend to create digital log models from the veneer sheet knot data for use in subsequent sawing simulations. The value to this approach is that the data collected on knot size and position can be used to investigate alternative processing methods.

For the larger veneer project involving acoustic velocities, we now have detailed information of knot size and counts as they vary within a tree from core to surface and from base to top. This will aid the development of improved models incorporating both acoustic and knot variables, for predicting stiffness from the whole tree through the value chain to the product.

Acknowledgements

References

Abbott, A.L., Kline, D.E., 2003. Imaging and analysis of logs and lumber. Comput. Electron. Agric. 41 (1–3), 1–6.

Amishev, D., Murphy, G.E., 2008. In-forest assessment of veneer grade Douglas-fir logs based on acoustic measurement of wood stiffness. For. Prod. J. 58 (11), 42–47.

Briggs, D.G., Thienel, G., Turnblom, E.C., Lowell, E., Dykstra, D., Ross, R., Wang, X., Carter, P., 2008. Influence of thinning on acoustic velocity of Douglas-fir trees in Western Washington and Western Oregon. In: Proceedings of the 15th International Symposium on Nondestructive Testing of Wood, Duluth, MN, Sept. 10–12, 2007, pp. 113–123.

Brunner, C.C., Maristany, A.G., Butler, D.A., Van Leeuwen, D., Funck, J.W., 1992. An evaluation of color spaces for detecting defects in Douglas-fir veneer. Ind. Metrol. 2 (3–4), 169–184.

Buehlmann, U., Lihra, T., Rancourt, V., Ait-Kadi, D., 2007. Detection capabilities of automated hardwood lumber defect-detection systems. For. Prod. J. 57 (10), 51–57.

Carter, P., Chauhan, S., Walker, J., 2006. Sorting logs and lumber for stiffness using director HM200. Wood Fiber Sci. 38 (1), 49–54.

Chauhan, S.S., Walker, J.C.F., 2006. Variations in acoustic velocity and density with age, and their interrelationships in radiata pine. For. Ecol. Manage. 229 (1–3), 388–394.

Cown, D., 2005. Understanding and managing wood quality for improving product value in New Zealand. N. Z. J. For. Sci. 35 (2–3), 205–220.

Dickson, R.L., Joe, B., Harris, P., Holtorf, S., Wilkinson, C., 2004. Acoustic segregation of Australian-grown Pinus radiata logs for structural board production. Aust. Forestry 67 (4), 261–266.

Fawcett, D., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. (27), 861–874.

Forrer, J.B., Butler, D.A., Funck, J.W., Brunner, C.C., 1988. Image sweep-and-mark algorithms. Part 1. Basic algorithms. For. Prod. J. 38 (11/12), 75–79.

Forrer, J.B., Butler, D.A., Brunner, C.C., Funck, J.W., 1989. Image sweep-and-mark algorithms. Part 2. Performance evaluations. For. Prod. J. 39 (1), 39–42.

Funck, J.W., Zhong, Y., Butler, D.A., Brunner, C.C., Forrer, J.B., 2003. Image segmentation algorithms applied to wood defect detection. Comput. Electron. Agric. 41 (1–3), 157–179.

Gonzalez, R.C., Woods, R.E., Eddins, S.L., 2004. Digital Image Processing Using MATLAB®. Prentice Hall, ISBN-13: 978-0130085191.

Huber, H.A., McMillin, C.W., McKinney, J.P., 1985. Lumber defect detection abilities of furniture rough mill employees. For. Prod. J. 35 (11/12), 79–82.

Kohavi, R., Provost, F., 1998. Glossary of terms. Editorial for the special issue on application of machine learning and the knowledge of discovery process. Mach. Learn. 30, 271–274.

Kumar, S., Jayawickrama, K.J.S., Lee, J., Lausberg, M., 2002. Direct and indirect measures of stiffness and strength show high heritability in a wind-pollinated radiata pine progeny test in New Zealand. Silvae Genet. 51 (5–6), 256–261.

Lasserre, J.-P., Mason, E.G., Watt, M.S., 2007. Assessing corewood acoustic velocity and modulus of elasticity with two impact based instruments in 11-year-old trees from a clonal-spacing experiment of Pinus radiata D. Don. For. Ecol. Manage. 239 (1–3), 217–221.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man. Cybern. SMC-9 (1), 62–66.

Polzleitner, W., Schwingshakl, G., 1992. Real-time surface grading of profiled wooden boards. Ind. Metrol. 2, 283–298.

Raymond, C.A., Joe, B., Anderson, D.W., Watt, D.J., 2008. Effect of thinning on relationships between three measures of wood stiffness in *Pinus radiata*: standing trees vs. logs vs. short clear specimens. Can. J. For. Res. 38 (11), 2870–2879.

Rinnhofer, A., Petutschnigg, A., Andreu, J.-P., 2003. Internal log scanning for optimizing breakdown. Comput. Electron. Agric. 41 (1–3), 7–21.

Ruz, G.A., Estevez, P.A., Perez, C.A., 2005. A neurofuzzy color image segmentation method for wood surface defect detection. For. Prod. J. 55 (4), 52–58.

Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. Inform. Process. Manage. 45 (4), 427–437.